The Fourth Transmembrane Segment of the Dopamine D2 Receptor: Accessibility in the Binding-Site Crevice and Position in the Transmembrane Bundle†

Jonathan A. Javitch,*,‡,§,⊥ Lei Shi,‡,§ Merrill M. Simpson,‡ Jiayun Chen,‡ Victor Chiappa,‡ Irache Visiers,|| Harel Weinstein,|| and Juan A. Ballesteros||,#

Center for Molecular Recognition, Departments of Pharmacology and Psychiatry, College of Physicians and Surgeons, Columbia University, 630 West 168th Street, New York, New York 10032, and Department of Physiology and Biophysics, Mount Sinai School of Medicine, New York, New York 10029

Received May 10, 2000; Revised Manuscript Received August 1, 2000

ABSTRACT: The binding site of the dopamine D2 receptor, like that of homologous G-protein-coupled receptors (GPCRs), is contained within a water-accessible crevice formed among its seven transmembrane segments (TMSs). Using the substituted-cysteine-accessibility method (SCAM), we are mapping the residues that contribute to the surface of this binding-site crevice. We have mutated to cysteine, one at a time, 21 consecutive residues in the fourth TMS (TM4). Eleven of these mutants reacted with charged sulfhydryl-specific reagents, and bound antagonist protected nine of these from reaction. For the mutants in which cysteine was substituted for residues in the cytoplasmic half of TM4, treatment with the reagents had no effect on binding, consistent with these residues being inaccessible and with the low-resolution structure of the homologous rhodopsin, in which TM3 and TM5 occlude the cytoplasmic half of TM4. Although hydrophobicity analysis positions the C-terminus of TM4 at 4.64, Pro−Pro and Pro−X−Pro motifs, which are known to disrupt α-helices, occur at position 4.59 in a number of homologous GPCRs. The SCAM data were consistent with a C-terminus at 4.58, but it is also possible that the α-helix extends one additional turn to 4.62 in the D2 receptor, which has a single Pro at 4.59. In homologous GPCRs, the high degree of sequence variation between 4.59 and 4.68 is more characteristic of a loop domain than a helical segment. This region is shown here to be very conserved within functionally related receptors, suggesting an important functional role for this putative nonhelical domain. This inference is supported by observed ligand-specific effects of mutations in this region and by the predicted spatial proximity of this segment to known ligand binding sites in other TMs.

The dopamine receptors, like the homologous receptors for the other biogenic amines, bind neurotransmitters present in the extracellular medium and couple this binding to the activation of intracellular G-proteins (*1*). The binding sites of these receptors are formed among their seven, mostly hydrophobic, transmembrane segments (TMSs)[1] (*2*) and are accessible to charged, water-soluble agonists, like dopamine. Thus, for each of these receptors, the binding site is contained within a water-accessible crevice, the binding-site crevice, extending from the extracellular surface of the receptor into the transmembrane domain. The surface of this crevice is formed by residues that can contact specific agonists and/or antagonists and by other residues that may play a structural role and affect binding indirectly.

Despite enormous efforts, no high-resolution structure of any G-protein coupled receptor (GPCR) has been determined experimentally. However, low-resolution structures that include a 9-Å projection structure of bovine rhodopsin (*3*) and a 7.5-Å structure of frog rhodopsin (*4*) have been used to guide the development of a molecular model of the TMSs of rhodopsin (*5*). Molecular models for many other GPCRs have been based on this model and on inferences from sequence alignments, analyzed in terms of conservation and physicochemical properties. These models provide a structural context, shared by all rhodopsin-like GPCRs, for incorporating constraints derived from biophysical and mutagenesis experiments into mechanistic hypotheses (*6*). Because mutagenesis experiments identified key ligand contact residues mainly in TM3, TM5, TM6, and TM7, and because in the initial projection map for rhodopsin TM4 appeared somewhat remote from the binding site, the fourth TMS has not received extensive consideration as a contributor to ligand binding. Nonetheless, on the basis of site-directed mutagenesis experiments, a number of laboratories have implicated residues in TM4 of GPCRs in the binding

† This work was supported by NIH Grants MH57324, MH54137, DA09083, DA00060; by the G. Harold and Leila Y. Mathers Charitable Trust; and by the Lebovitz Trust.

* To whom correspondence should be addressed. Phone: 212-305-7308. Fax: 212-305-5594. E-mail: jaj2@columbia.edu.

‡ Center for Molecular Recognition.

§ Department of Pharmacology.

⊥ Department of Psychiatry.

|| Mount Sinai School of Medicine.

Present address: Novasite Pharmaceuticals, Inc., 3520 Dunhill Street, San Diego, CA 92121.

[1] Abbreviations: GPCRs, G-protein-coupled receptors; TMS, transmembrane segment; TMX, the Xth transmembrane segment; SCAM, substituted-cysteine accessibility method; MTS, methanethiosulfonate; MTSEA, MTSethylammonium; MTSET, MTSethyltrimethylammonium; MTSES, MTSethylsulfonate; SS, sequence similarity index; NSI, normalized similarity index; NT, N-terminus; CT, C-terminus; PP, Pro−Pro; PXP, Pro−X−Pro.

10.1021/bi001069m CCC: $19.00
Published on Web 09/13/2000

of agonists and antagonists (reviewed in ref *7*), although it is not clear whether the effects of these mutations on binding result from alterations of direct interaction or from indirect effects. In a more direct demonstration of the proximity of TM4 residues to the ligand binding site, an agonist and an antagonist photoaffinity ligand were found to label covalently part of TM4 in the α2 adrenergic receptor, although the specific residue labeled could not be identified (*8*).

Here we report the application of the substituted-cysteine accessibility method (SCAM) to identify systematically all the residues in TM4 of the D2 receptor that contribute to the binding-site crevice. We have been using SCAM (*9*) to identify the residues in TM2, TM3, TM5, TM6, and TM7 that form the surface of the binding-site crevice in the human dopamine D2 receptor (*10*−*15*). Consecutive residues in the TMSs are mutated to cysteine, one at a time, and the mutant receptors are expressed in heterologous cells. If ligand binding to a cysteine-substitution mutant is near-normal, we assume that the structure of the mutant receptor, especially around the binding site, is similar to that of wild type and that the substituted cysteine lies in a similar orientation to that of the wild-type residue. In the TMSs, the sulfhydryl of a cysteine facing into the binding-site crevice should react much faster with charged sulfhydryl-specific reagents than should sulfhydryls facing into the interior of the protein or into the lipid bilayer. For such reagents, we use derivatives of methanethiosulfonate (MTS): positively charged MTSethylammonium (MTSEA) and MTSethyltrimethylammonium (MTSET) and negatively charged MTSethylsulfonate (MTSES) (*16*). These reagents are about the same size as dopamine, with maximum dimensions of approximately 10 by 6 Å. They form mixed disulfides with the cysteine sulfhydryl, covalently linking -SCH_2CH_2X, where X is NH_3^+, $N(CH_3)_3^+$, or SO_3^-. We use two criteria for identifying an engineered-cysteine as forming the surface of the binding-site crevice: (i) The reaction with an MTS reagent alters binding irreversibly. (ii) This reaction is retarded by the presence of ligand.

The SCAM analysis is presented here in the context of results from sequence analysis and molecular modeling of TM4. This has allowed us to consider our empirical findings regarding the accessibility of the residues in TM4 in the structural context of complementary efforts to develop and refine computational molecular models of GPCRs.

EXPERIMENTAL PROCEDURES

Numbering of Residues. Residues are numbered according to their positions in the human dopamine D2 receptor sequence. In some cases, we also index residues relative to the most conserved residue in the TMS in which it is located (*6*). By definition, the most conserved residue is assigned the position index "50", e.g., $W160^{4.50}$, and therefore $V159^{4.49}$ and $V161^{4.51}$. This indexing simplifies the identification of aligned residues in different GPCRs.

Site-Directed Mutagenesis. Cysteine mutations were generated as described previously (*11*). Mutations were confirmed by DNA sequencing. Mutants are named as (wild-type residue)(residue number)$^{\text{position index}}$(mutant residue), where the residues are given in the single-letter code.

Transfection. The cDNA encoding the dopamine $D2_{short}$ receptor or the appropriate cysteine mutant, epitope tagged at the amino terminus with the cleavable influenza-hemagglutinin signal sequence followed by the "FLAG" epitope (IBI, New Haven, CT) (*14*) in the bicistronic expression vector pcin4 (a gift from Dr. S. Rees, Glaxo) (*17*) was used for all transfections, which were performed as described previously (*14*).

HEK 293 cells in DMEM/F12 (1:1) with 10% bovine calf serum (Hyclone) and 293-TSA cells (a clonal line of HEK 293 cells stably expressing the SV40 large T antigen) in DMEM with 10% fetal calf serum were maintained at 37 °C and 5% CO_2. For transient transfection, 35-mm dishes of 293-TSA cells at 70−80% confluence were transfected with 2 μg of wild type or mutant D2 receptor cDNA in pcin4 (see above) using 9 μL of lipofectamine (Gibco) and 1 mL of OPTIMEM (Gibco). Five hours after transfection, the solution was removed and fresh media was added. Twenty-four hours after transfection the media was changed. Forty-eight hours after transfection, cells were harvested as described below. For stable transfection, HEK 293 cells were transfected with D2 receptor cDNA in pcin4 as described above. Twenty-four hours after transfection the cells were split to a 100-mm dish, and 700 μg/mL Geneticin was added to select for a stably transfected pool of cells.

Harvesting Cells. Cells were washed with phosphate-buffered saline (PBS; 8.1 mM NaH_2PO_4, 1.5 mM KH_2PO_4, 138 mM NaCl, and 2.7 mM KCl, pH 7.2), briefly treated with PBS containing 1 mM EDTA, and then dissociated in PBS. Cells were pelleted at 1000*g* for 5 min at 4 °C and resuspended for binding or treatment with MTS reagents.

[^{3}H]N-Methylspiperone Binding. Whole cells from a 35-mm plate were suspended by pipetting in 400 μL of buffer A (25 mM HEPES, 140 mM NaCl, 5.4 mM KCl, 1 mM EDTA, and 0.006% BSA, pH 7.4). Cells were then diluted with buffer A, typically 20-fold. Depending on the level of expression in the various mutants, adjustments in the number of cells per assay tube were made as necessary to prevent depletion of ligand in the case of very high expression or to increase the signal in the case of low expression. [^{3}H]*N*-Methylspiperone (Dupont/NEN) binding was performed as described previously (*13*).

Reactions with MTS Reagents. Whole cells from a 35-mm plate were suspended in 400 μL buffer A. Aliquots (45 μL) of cell suspension were incubated with freshly prepared MTS reagents (5 μL) at the stated concentrations at room temperature for 2 min. Cell suspensions were then diluted 16-fold, and 200 μL aliquots were used to assay for [^{3}H]*N*-methylspiperone (200 pM) binding as described (*13*). The fractional inhibition was calculated as 1 − [(specific binding after MTS reagent)/(specific binding without reagent)]. We used SPSS for Windows (SPSS, Inc.) to analyze the effects of the MTS reagents by one-way ANOVA with Dunnett's post hoc test ($p < 0.05$).

The second-order rate constant (k) for the reaction of MTSEA with each susceptible mutant was estimated by determining the extent of reaction after a fixed time, 2 min, with six concentrations of MTSEA (typically 0.025−10 mM) (all in excess over the reactive sulfhydryls). The fraction of initial binding, Y, was fit to (1-plateau)e^{-kct} + plateau, where plateau is the fraction of residual binding at saturating concentrations of MTSEA, k is the second-order rate constant (in M^{-1} s^{-1}), c is the concentration of MTSEA (M), and t is the time (120 s).

Table 1: Characteristics of [^{3}H]*N*-Methylspiperone Binding to Cysteine-Substituted Dopamine D2 Receptor[a]

mutant	K_D (pM)	$K_{MUT}/K_{C118^{3.36}S}$	B_{MAX} (fmol/cm^2)	*n*
V154$^{4.44}$C	70 ± 4	0.8	317 ± 4	2
M155$^{4.45}$C	60 ± 4	0.7	228 ± 23	2
I156$^{4.46}$C	53 ± 10	0.6	214 ± 37	2
S157$^{4.47}$C	105 ± 7	1.2	418 ± 14	3
I158$^{4.48}$C	108 ± 18	1.3	283 ± 75	3
V159$^{4.49}$C	45 ± 5	0.5	137 ± 24	2
W160$^{4.50}$C	42 ± 6	0.5	62 ± 26	2
V161$^{4.51}$C	87 ± 12	1.0	306 ± 19	2
L162$^{4.52}$C	97 ± 24	1.1	256 ± 51	3
S163$^{4.53}$C	94 ± 10	1.1	90 ± 12	2
F164$^{4.54}$C	63 ± 2	0.7	208 ± 14	2
T165$^{4.55}$C	80 ± 16	0.9	422 ± 73	2
I166$^{4.56}$C	81 ± 10	1.0	260 ± 81	2
S167$^{4.57}$C	87 ± 16	1.0	323 ± 39	2
C168$^{4.58}$ (C118$^{3.36}$S)	85 ± 4	1.0	320 ± 72	2
P169$^{4.59}$C	134 ± 20	1.6	228 ± 19	2
L170$^{4.60}$C	74 ± 4	0.9	112 ± 20	2
L171$^{4.61}$C	58 ± 3	0.7	57 ± 18	2
F172$^{4.62}$C	97 ± 3	1.1	63 ± 7	2
G173$^{4.63}$C	61 ± 13	0.7	99 ± 15	2
L174$^{4.64}$C	73 ± 3	0.9	199 ± 29	4
N175$^{4.65}$C	114 ± 24	1.3	204 ± 39	3

[a] Cells transfected with the appropriate receptor were assayed as described in Experimental Procedures. Data were fit to the binding isotherm by nonlinear regression. The means and SEM are shown for *n* independent experiments, each with duplicate determinations. B_{MAX} values are presented as femtomoles per square centimeter of plate area. Note that C118$^{3.36}$S, the background for all the mutations, contains Cys168, which is present in all the mutants.

Table 2: Inhibitory Potency of (−)-Sulpiride on [^{3}H]*N*-methylspiperone Binding to Cysteine-Substituted Dopamine D2 Receptor[a]

mutant	apparent K_I (nM)	$K_{I(MUT)}/K_{I(C118^{3.36}S)}$	*n*
V154$^{4.44}$C	18 ± 2	1.1	2
M155$^{4.45}$C	16 ± 2	1.0	2
I156$^{4.46}$C	19 ± 2	1.2	2
S157$^{4.47}$C	13 ± 1	0.8	2
I158$^{4.48}$C	16 ± 1	1.0	3
V159$^{4.49}$C	21 ± 0	1.3	2
W160$^{4.50}$C	156 ± 20	9.8	2
V161$^{4.51}$C	22 ± 4	1.4	2
L162$^{4.52}$C	17 ± 2	1.1	2
S163$^{4.53}$C	28 ± 2	1.7	3
F164$^{4.54}$C	34 ± 4	2.1	2
T165$^{4.55}$C	17 ± 5	1.0	3
I166$^{4.56}$C	35 ± 1	2.2	2
S167$^{4.57}$C	21 ± 1	1.3	2
C168$^{4.58}$ (C118$^{3.36}$S)	16 ± 4	1.0	2
P169$^{4.59}$C	3737 ± 566	234.3	2
L170$^{4.60}$C	70 ± 19	4.4	2
L171$^{4.61}$C	97 ± 15	6.1	2
F172$^{4.62}$C	298 ± 97	18.7	3
G173$^{4.63}$C	99 ± 11	6.2	3
L174$^{4.64}$C	87 ± 31	5.5	5
N175$^{4.65}$C	31 ± 7	1.9	3

[a] Cells transfected with the appropriate receptor were assayed with [^{3}H]*N*-methylspiperone (150 pM) as described in Experimental Procedures in the presence of nine concentrations of (−)-sulpiride. The apparent K_I was determined by the method of Cheng and Prusoff (*51*) using the IC_{50} value obtained by fitting the data to a one-site competition model by nonlinear regression. The means and SEM are shown for *n* independent experiments, each with duplicate determinations. Note that C118$^{3.36}$S, the background for all the mutations, contains Cys168, which is present in all the mutants.

Analysis of Sequence Conservation. The sequence similarity index (SS) is derived from sequence variability (*18*). SS makes use of the exchange matrix, and SS at one position is defined as the weighted sums of residue similarities over all sequence pairs within an alignment. Data for SS were mostly retrieved from the corresponding profiles in GPCRDB (*19*). The profiles for the dopamine D2-like (D2, D3, D4) receptor sequences and for the D2 receptor sequences were calculated using the WHAT IF program (*20*) using the same methods as described for the GPCRDB.

The SS of region 4.59−4.68, region 1.21−1.30, and all TMS residues were averaged. The conservation of region 4.59−4.68 was normalized [the normalized similarity index (NSI)] as the following: (avg 4.59−4.68 − avg 1.21−1.30)/(avg TMSs − avg 1.21−1.30). Therefore, the NSI describes the degree of conservation of region 4.59−4.68 as compared to that of region 1.21−1.30 relative to the degree of conservation of the TMSs. Comparisons were made between different receptors and receptor subtypes within the amine family of rhodopsin-like GPCRs.

Comparative Molecular Modeling. The 3D model of the TMSs of the D2 receptor uses the helix axes of the template described by Baldwin (*5*) following the 7.5 Å rhodopsin electronmicroscopy map (*4*). The transmembrane bundle incorporates the predictions from the criteria discussed in Results for the identification of the transmembrane span of the TMSs, and their inward/outward orientation, as well as experimental constraints of homologous GPCRs that include information about ligand interaction at positions D3.32, S5.43, S5.46, and F6.52 (reviewed in ref *7*), the role and position of the arginine cage at the cytoplasmic end of TM3 (*21*), and the specific structural information from SCAM of TM2, TM3, TM5, TM6, and TM7 (*11*−*14*, *22*). Criteria used for the vertical positioning of TMSs relate to the high concentration of nonconserved Arg and Lys residues at the cytoplasmic ends of TMSs (*23*, *24*), as well as identified proximities between TMSs (*25*, *26*).

RESULTS

Application of the Substituted-Cysteine Accessibility Method to TM4 of the Dopamine D2 Receptor. (1) Effects of Cysteine-Substitution on Antagonist Binding. In a background of the mutant C118$^{3.36}$S, which is relatively insensitive to the MTS reagents (*10*), we mutated to cysteine, one at a time, 21 consecutive residues, V154$^{4.44}$ to N175$^{4.65}$, in TM4. Each mutant receptor was transiently or stably expressed in HEK 293 cells, and the K_D and B_{MAX} characterizing the equilibrium binding of the radiolabeled antagonist, [^{3}H]*N*-methylspiperone, were determined. At all 21 positions, the K_D of the cysteine-substitution mutant was between 0.5 and 1.6 times the K_D of C118S and the B_{MAX} ranged from 18 to 130% of that obtained with C118S (Table 1).

The K_I of the antagonist sulpiride in competition with [^{3}H]-*N*-methylspiperone was determined in the 21 mutants (Table 2). At 21 positions, the K_I was between 0.8 and 2.2 times the K_I of C118S. In the four mutants L170C, L171C, G173C, and L174C, the K_I was between 4.4 and 6.2 times the K_I of C118S, whereas in W160C, P169C, and F172C, the K_I was 10, 234, and 19 times, respectively, the K_I of C118S.

(2) Reactions of Charged MTS Reagents with the Mutants. A 2-min application of 2.5 mM MTSEA significantly inhibited [^{3}H]*N*-methylspiperone binding to 11 of 21 cys-

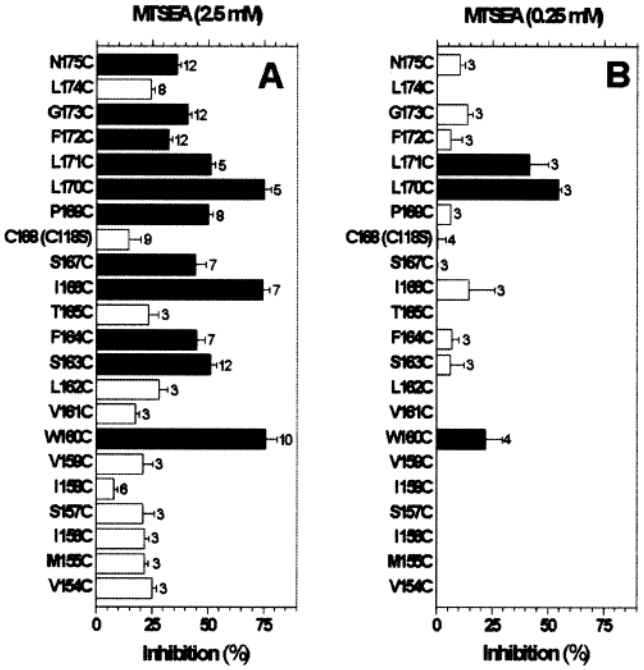

FIGURE 1: Effects of MTSEA application on specific [^{3}H]*N*-methylspiperone (200 pM) binding to intact cells transfected with wild-type or mutant D2 receptors. Cells were treated with (A) 2.5 mM MTSEA or (B) 0.25 mM MTSEA for 2 min. The means and SEM are shown. The number of independent experiments for each mutant is shown next to the bars. Solid bars indicate mutants for which inhibition was significantly different ($p < 0.05$) than C118$^{3.36}$S by one-way ANOVA. Note that C118$^{3.36}$S, the background for all the mutations, contains Cys168, which is present in all the mutants. Only the mutants that were significantly different from C118$^{3.36}$S were studied at 0.25 mM in panel B.

Table 3: Rates of Reaction of MTSEA with Cysteine-Substituted Dopamine D2 Receptor[a]

mutant	k_{MTSEA} (M^{-1} s^{-1})	k_{MUT}/k_{WT}	*n*
W160$^{4.50}$C	9 ± 2	0.2	4
S163$^{4.53}$C	3 ± 1	0.1	3
F164$^{4.54}$C	4 ± 1	0.1	3
I166$^{4.56}$C	7 ± 2	0.2	3
S167$^{4.57}$C	2 ± 1	0.05	3
P169$^{4.59}$C	5 ± 1	0.1	3
L170$^{4.60}$C	74 ± 15	1.9	3
L171$^{4.61}$C	43 ± 12	1.1	3
F172$^{4.62}$C	5 ± 1	0.1	3
G173$^{4.63}$C	5 ± 1	0.1	3
N175$^{4.65}$C	4 ± 1	0.1	3

[a] The second-order rate constant (*k*) was determined as described in Experimental Procedures. The means and SEM of *n* independent experiments, each performed with triplicate determinations, are shown. k_{MUT}/k_{WT} was obtained by dividing each *k* value by the *k* determined for the wild-type receptor in which Cys118 reacts, 39 M^{-1} s^{-1} (*11*), which is not present in the background used to construct the Cys substitution mutants.

teine-substitution mutants (Figure 1A). A 2-min application of 0.25 mM MTSEA significantly inhibited binding to 3 of these 11 mutants (Figure 1B). To quantitate the susceptibility to MTSEA, we determined the second-order rate constants for the reaction with MTSEA (Table 3). The most reactive cysteines were those substituted for L170 and L171. Cysteines substituted for W160 and I166 were of intermediate reactivity with MTSEA. Cysteines substituted for S163, F164, S167, P169, F172, G173, and N175 were least reactive.

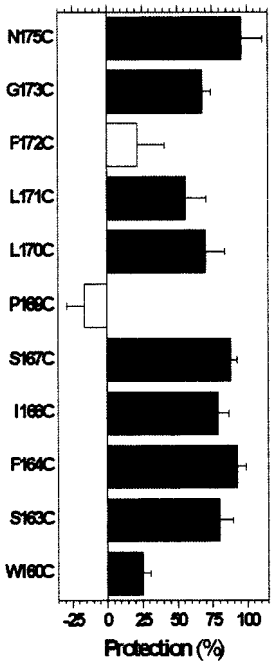

FIGURE 2: Sulpiride protection of cysteine-substitution mutants. Dissociated cells were incubated in buffer A for 20 min at room temperature in the presence or absence of (±)-sulpiride, and then MTSEA was added, in the continued presence or absence of sulpiride, for 2 min at a concentration chosen to inhibit 50−75% of specific binding in the absence of sulpiride. Concentrations of MTSEA were as follows: 5 mM: F164C; 2.5 mM: S163C, S167C, P169C, F172C, G173C, N175C; 1 mM: W160C, I166C; 0.25 mM or 0.1 mM: L170C, L171C. For 5 of the 11 mutants, S163C, F164C, I166C, S167C, and N175C sulpiride was used at a concentration of 10 μM. To compensate for changes in the K_I, increased concentrations of sulpiride were used for the following mutants: 50 μM: L170C, L171C, G173C; 100 μM: W160C, F172C; 400 μM: P169C. Protection was calculated as 1− [(inhibition in the presence of sulpiride)/(inhibition in the absence of sulpiride)]. Solid bars indicate that protection by sulpiride was significant ($p < 0.05$) by paired t-test.

The reversible antagonist sulpiride significantly retarded the reaction of MTSEA with W160C, S163C, F164C, I166C, S167C, L170C, L171C, G173C, and N175C. The degree of protection varied from 25 to 96% (Figure 2). Although sulpiride slightly retarded the reaction of MTSEA with F172C, the protection was not significant. Sulpiride had no effect on the reactivity of P169C.

Nine of the mutants, W160C, F164C, I166C, P169C, L170C, L171C, F172C, G173C, and N175C, were susceptible to reaction with 1 mM MTSET (Figure 3A). At 10 mM, MTSES, the negatively charged derivative, significantly inhibited binding to five of these nine mutants, W160C, L170C, L171C, G173C, and N175C (Figure 3B).

A Structural Template for TM4 of the Dopamine D2 Receptor. (1) Prediction of Helix Ends and Orientation Based on Sequence Analysis. A hydrophobicity plot, in which the average hydrophobicity is plotted against the D2 TM4 receptor sequence, identifies the TM4 segment from V4.44 to L4.64 as a continuous stretch with a high overall hydrophobicity. However, the predicted N-terminus (NT) at 4.44 is one turn above the end of the TM4 helix in rhodopsin that was determined experimentally to be at position 4.40 (*27*). Analysis of the helical periodicity of the conservation pattern by Fourier transform of the variability provides an alternative algorithm for predicting the boundaries of the

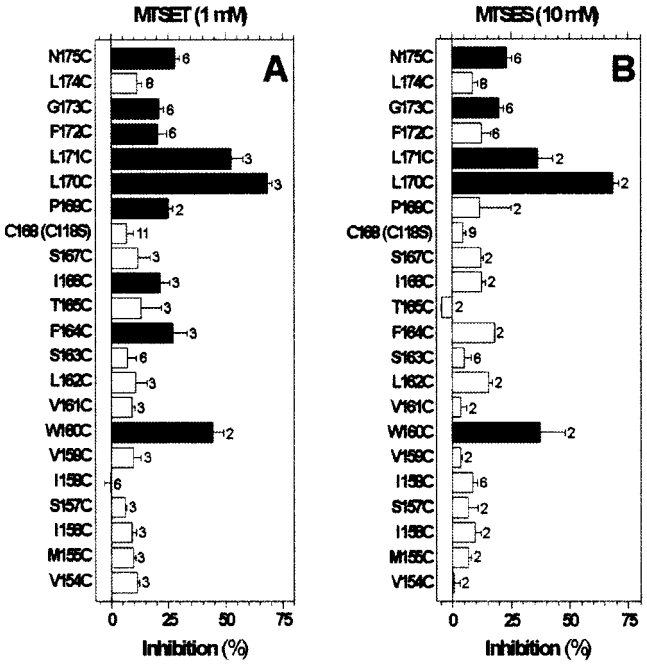

FIGURE 3: Effects of MTSET and MTSES application on specific [^{3}H]*N*-methylspiperone (200 pM) binding to intact cells transfected with wild-type or mutant D2 receptors. Cells were treated with (A) 1 mM MTSET or (B) 10 mM MTSES for 2 min. The inhibition of specific [^{3}H]*N*-methylspiperone (200 pM) binding to intact cells transfected with wild-type or mutant D2 receptors resulting from a 2-min application of (A) 1 mM MTSET or (B) 10 mM MTSES. On the basis of the relative rate constants for reaction with simple thiols in solution, namely, 10:4:1 for MTSET, MTSEA, and MTSES, (*16*), we used equireactive concentrations of 1 mM MTSET (A), 2.5 mM MTSEA (Figure 1A), and 10 mM MTSES (B). The means and SEM are shown. The number of independent experiments for each mutant is shown next to the bars. Solid bars indicate mutants for which inhibition was significantly different ($p < 0.05$) than C118$^{3.36}$S by one-way ANOVA. Note that C118$^{3.36}$S, the background for all the mutations, contains Cys168, which is present in all the mutants.

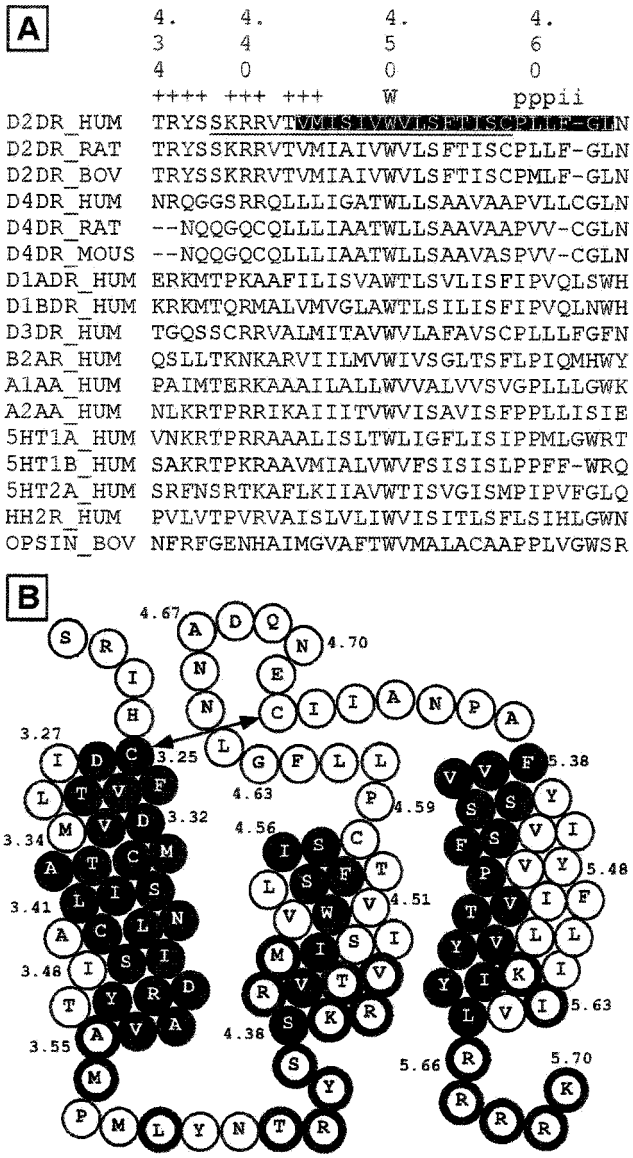

FIGURE 4: Sequence analysis of TM4 of the dopamine D2 and homologous neurotransmitter receptors. In panel A, an alignment of the D2 with representative homologous neurotransmitter receptors is shown, indicating the presence of nonconserved Arg and Lys (indicated by a "+" over the alignment). The presence of prolines in an aligned sequence is indicated by a "p" over the alignment, and the presence of insertions is indicated by an "i". The residues in the human D2 receptor predicted to be in the TM4 α-helix (see text) are underlined. Residues predicted to reside in the membrane based on a hydrophobicity plot are shaded (see text). In panel B, a helical net representation of TM3, TM4, and TM5 of the D2 receptor is shown, with the conserved residues predicted to face into the binding-site crevice or the protein interior shaded. The positions in which nonconserved Arg and Lys appear in homologous receptors are shown with a thick circle.

TMSs (*28*−*30*). The α-helix periodicity index calculated for TM4 of neurotransmitter GPCRs from the Fourier transform of the variability (number of different amino acid at a given position) suggests that the NT of TM4 is positioned at 4.38 (see Figure 8C in ref *6* and discussion therein), more in line with the experimental data in rhodopsin.

In known transmembrane helical structures, nonconserved Arg and Lys residues tend to concentrate at the cytoplasmic boundary of TM helices where they face lipid at the level of the negatively charged phospholipid headgroups, anchoring the TM helix to the membrane through ionic interactions (*23*). A multiple sequence alignment of 62 neurotransmitter GPCRs homologous to the D2 receptor was constructed, and the presence of nonconserved Arg and Lys residues within TM4 was identified (Figure 4A). From 4.34 to 4.45, every position except 4.38 and 4.42 contains Arg or Lys residues in some receptors. Arg/Lys in the TM4 α-helix at the cytoplasmic interface should concentrate on one face of the α-helix, facing the phospholipid headgroups. A helical net representation of TM4 (Figure 4B) shows that the pattern of Arg/Lys present across these 62 GPCRs is consistent with α-helical character from 4.38 until 4.46. This periodicity is not maintained before 4.38. Thus, although a simple hydrophobicity plot excludes the residues before 4.44 from the helix, the Arg/Lys criteria suggest that the TM4 helix continues to the highly conserved 4.38.

The pattern of amino acid conservation is also consistent with S4.38 as the NT for TM4. Conserved residues are predicted to face the protein interior, whereas nonconserved residues such as the Arg/Lys motif described above are predicted to face the solvent, in this case the lipid milieu. S4.38 and I4.46 are highly conserved in amino acid character and volume and are predicted, therefore, to face inside (shaded in Figure 4B). Consistent with these predictions are

Table 4: Sequence Similarity Analysis of the Extracellular Portion of TM4[a]

		amine receptors	dopamine receptors	dopamine D1-like (D1, D5) receptors	dopamine D2-like (D2, D3, D4) receptors	dopamine D2 receptors	adrenergic receptor	β-adrenergic receptors	β2-adrenergic receptors	serotonin receptors	serotonin 5HT2 receptors
no. of seq		210	43	22	13	6	60	25	8	65	12
position in region	4.59	P 58	P 50	I 95	P 100	P 100	P 53	L 75	L 100	P 72	P 100
	4.60	P 68	P 58	P 100	L 70	L 83	P 100	P 100	P 100	P 78	I 90
	4.61	L 64	V 70	V 100	L 86	L 100	L 70	I 100	I 100	L 54	P 100
	4.62	L 55	Q 51	Q 100	F 79	F 100	I 59	M 44	Q 88	F 70	V 90
	4.63	G 31	G 56	L 100	G 100	G 100	S 44	M 67	M 100	W 38	F 53
	4.64	W 33	L 30	N 59	L 71	L 100	H 30	H 71	H 100	G 37	G 100
	4.65	W 33	N 51	W 100	N 100	N 100	W 41	W 100	W 100	Q 34	L 75
	4.66	H 27	H 56	H 100	N 58	N 100	W 40	W 80	Y 100	A 29	Q 70
	4.67	K 37	K 48	K 86	T 70	T 70	R 50	R 96	R 88	K 42	D 75
	4.68	A 42	A 40	A 68	D 45	D 67	A 50	A 54	A 100	A 38	D 90
avg	4.59−4.68	45 ± 15	51 ± 11	91 ± 15	78 ± 19	92 ± 13	54 ± 20	79 ± 20	98 ± 5	49 ± 18	84 ± 16
avg	1.21−1.30	30 ± 5	41 ± 6	48 ± 14	53 ± 12	72 ± 8	38 ± 7	44 ± 9	85 ± 11	33 ± 6	56 ± 12
avg TMSs		68 ± 19	76 ± 19	93 ± 11	86 ± 16	98 ± 7	74 ± 20	86 ± 17	98 ± 6	70 ± 19	89 ± 16
normalized similarity index		0.39	0.29	0.95	0.76	0.78	0.43	0.84	0.96	0.44	0.86

[a] For each group of receptors at a given index position in region 4.59−4.68, the most common residue at this position, and the sequence similarity (SS) are shown, e.g., a 100% conserved position is characterized with an SS of 100. Avg 1.21−1.30 represents the average SS of residues 1.21−1.30, avg 4.59−4.68 represents the average SS of residues 4.59−4.68, and average TMSs represents the average SS ± SD of all the TMSs residues. The normalized similarity index (NSI) is calculated as described in Experimental Procedures.

spin labeling studies of rhodopsin that identified position 4.42 as facing the protein interior, whereas positions 4.40, 4.41, 4.43, and 4.44 faced the lipid milieu (*27*).

W4.50 is completely conserved, and positions 4.53, 4.56, and 4.57 are highly conserved in volume and/or hydrogen bonding character. Thus, all of these residues are predicted to face the protein interior (Figure 4B). Position 4.54 is less conserved but shows a subtype-selective pattern, and is, therefore, also predicted to face inward. In contrast, the residues at 4.52 and 4.58 preserve a high average hydrophobicity but are nonconserved, an indication that they face lipid. Thus the pattern of conservation is also consistent with α-helical periodicity from S4.38 until C4.58 (Figure 4B).

The TM4 segment from 4.59 to 4.64 is highly hydrophobic and thus predicted to be in the membrane as illustrated schematically in Figure 4B. Consistent with this prediction, a highly hydrophobic photoaffinity probe labeled position 4.64 in rhodopsin (*31*). Fourier transform of the conservation pattern predicts 4.63 as the C-terminus (CT) of the TM4 α-helix. However, the presence of insertions at positions 4.62−4.63 (Figure 4A), the presence and positioning of Pro residues after position 4.58 (see below), and the high degree of sequence divergence (see below), make it unlikely that the region after 4.58 has a highly ordered secondary structure in all homologous receptors.

(2) Effects of Prolines near the Carboxy-Termini of α-Helices. The proposed span of TM4 in the dopamine receptor includes a Pro at position $169^{4.59}$. Many other GPCRs also contain a single Pro at position 4.59. Other receptors contain a single Pro at 4.60 or no Pro at all (Figure 4A). Other patterns frequently present in this region include Pro−Pro (PP) or Pro−X−Pro (PXP) motifs (Figure 4A), which are known to disrupt secondary structural elements and thus are almost never present within α-helices (*32*).

To explore the potential impact of the various Pro motifs on the secondary structure of TM4, we performed an Iditis (V3.31, Oxford Molecular Ltd.) search of the protein data bank (*33*) of all known high-resolution structures (≤2.0 Å) containing single P, PP, or PXP motifs near the C-terminal end of an α-helical stretch of at least eight residues. The most common structure was one in which a single Pro was at the site of termination of an α-helix (302 hits). A number of structures were identified, however, in which the α-helical secondary structure continued for two to four residues following a single Pro (13 hits). In contrast, no structure was found in which PP or PXP were present in an α-helical conformation, although this motif did occur immediately after the termination of preceding α-helix (seven hits for PP and 12 hits for PXP). Thus, the most likely CT of the helix is 4.58, although in receptors containing a single Pro at position 4.59 (or 4.60), such as the D2 receptor, the helix could potentially extend to 4.62 or 4.63.

It should be noted that given the scarcity of high resolution structures of membrane proteins, the results of the Iditis search are most pertinent to water-soluble proteins. For membrane-embedded proteins, the hydrophobic environment strongly favors regular secondary structures such as α-helices that satisfy the backbone H-bonding potential, and Pro-containing α-helices are common in membrane proteins such as receptors, transporters, and ion channels. Nonetheless, given the lack of high resolution structural information of membrane proteins, the data on the C-termini of helices containing Pro are the best guide currently available.

(3) Analysis of Sequence Conservation of the Extracellular Portion of TM4. In homologous amine GPCRs, the high degree of sequence variation observed at the extracellular portion of TM4 is more characteristic of a loop domain than a helical TMS. We observed, however, that this region appeared to be rather conserved within functionally related receptors as well as within species variants of the same receptor subtype. We quantitated this observation by analyzing the sequence conservation as described in Experimental Procedures, and the analysis is shown in Table 4.

Our analysis evaluated the sequence conservation in region 4.59−4.68 in comparison to that of the 10 residues in the

extracellular domain immediately preceding TM1 (region 1.21−1.30), which was chosen as an example of a region expected to be nonconserved. Table 4 shows the position index, the most common residue at this position, and the sequence similarity parameter (SS) of each residue in region 4.59−4.68 within the indicated groups and subgroups of amine receptors. Average SS values over each 10-residue stretch, and over all the residues in the TMSs, are shown below each group. The "normalized similarity index" (NSI) was calculated to quantify how similar the conservation of region 4.59−4.68 is to the overall conservation of the TMSs, as a standard of high conservation. Thus, 0.0 indicates no more conservation than region 1.21−1.30, and 1.0 indicates an extent of conservation equal to that of the average TMSs.

The results in Table 4 show that, although this extracellular portion of TM4 is not more conserved in all dopamine receptors (0.29) than in all amine receptors (0.39), this changes when dopamine receptors are divided into D1-like (D1, D5) or D2-like (D2, D3, D4) receptors. For these subgroups, the conservation is dramatically higher relative to the conservation of region 1.21−1.30 (0.95 and 0.76, respectively). Within the species variants of the D2 receptor alone, the NSI (0.78) was similar to that of the D2-like receptors, consistent with the structural similarity expected from the closely related pharmacology and function of these receptors. Analysis of the adrenergic receptors and serotonin receptors identified a similar pattern of high conservation of the extracellular portion of TM4 within functionally related subgroups of receptors, such as the β adrenergic receptors (Table 4).

DISCUSSION

Residues Forming the Water-Accessible Surface of the Binding-Site Crevice. We identify residues on the water-accessible surface of the D2 receptor by the ability of the MTS reagents to react with substituted cysteine residues. The MTS reagents react $>10^9$ times faster with ionized thiolates than with unionized thiols (*34*), and only water-accessible cysteines are likely to ionize to a significant extent. Moreover, MTSEA, MTSET, and MTSES are charged and hydrophilic. Thus, we assume that these reagents will react much faster with water-accessible cysteine residues than with cysteines facing the protein interior or lipid. On the basis of the near-normal affinities of the cysteine mutants for *N*-methylspiperone, it is likely that the global structures of the mutants are near-normal and that the substituted cysteines are reliable reporters for the accessibility of the wild-type residue for which they are substituted. We infer that the MTS reagents have reacted if the binding of ligand is irreversibly affected. Although the mechanism could be steric, electrostatic, or indirect, a change in binding is taken as evidence for reaction. Thus, we infer from the results that 11 of the 21 residues tested are on the water-accessible surface of the D2 receptor. These include W160$^{4.50}$, S163$^{4.53}$, F164$^{4.54}$, I166$^{4.56}$, S167$^{4.57}$, P169$^{4.59}$, L170$^{4.60}$, L171$^{4.61}$, F172$^{4.62}$, G173$^{4.63}$, and N175$^{4.65}$ (Figure 1).

The residues that form the surface of the binding-site crevice are a subset of the water-accessible residues. Water-accessible residues are considered to be in the binding-site crevice if competitive antagonists or agonists retard the reaction of the MTS reagents. The competitive antagonist sulpiride protected 9 of the 11 residues tested, although the extent of protection by sulpiride varied considerably among the mutants (Figure 2). From the protection data, we infer that P169$^{4.59}$ and F172$^{4.62}$ are likely located at the margin of the binding site such that the presence of sulpiride in the binding site decreases access of MTSEA only in part, resulting in no or relatively poor protection. Protection of a substituted cysteine is most simply explained by its proximity to the sulpiride-binding site. Nevertheless, not every one of the sites protected by sulpiride needs to make direct contact with the drug. Sulpiride could protect residues deeper in the crevice by binding above them and blocking the passage of MTSEA from the extracellular medium toward the cytoplasmic end of the crevice. We also cannot rule out indirect effects through propagated structural changes for protection by the antagonist sulpiride.

Given the differences in the binding of sulpiride and *N*-methylspiperone to particular mutants, it would be interesting to compare the abilities of *N*-methylspiperone and sulpiride to protect, particularly in P169C and F172C. Unfortunately, because *N*-methylspiperone is rather lipophilic, it is very difficult to remove it sufficiently to allow reliable measurement of residual binding. In contrast, sulpiride is hydrophilic and its binding is sodium-dependent, two factors that facilitate its washout. Nonetheless, we have observed that the addition of MTSEA to P169C and to F172C altered the rate of dissociation of prebound [^{3}H]*N*-methylspiperone (data not shown). This argues that these residues are not sterically protected by bound *N*-methylspiperone, consistent with the results of sulpiride protection experiments in these mutants.

Comparison of Reactivity with MTSEA, MTSET, and MTSES. MTSEA is mostly positively charged at physiological pH, whereas MTSET has a fixed positive charge. At physiological pH, the negatively charged MTSES is essentially completely ionized. MTSET significantly inhibited binding to all the mutants inhibited by 2.5 mM MTSEA, except S163$^{4.53}$C and S167$^{4.57}$C (Figure 3A), the two least reactive mutants (Table 3). In addition, the inhibition following 2 min of 10 mM MTSES did not reach statistical significance at F164$^{4.54}$C, I166$^{4.56}$C, P169$^{4.59}$C, and F172$^{4.62}$C mutants; the inhibition after 2 min of 1 mM MTSET was also not robust for these mutants (Figure 3). At W160$^{4.50}$C MTSET and MTSES produced about the same submaximal inhibition of binding, suggesting that the rates of reaction of the positively and negatively charged reagents are roughly similar and that the electrostatic potential at this position is near neutral. This contrasts with our findings in TM2 and TM3, where we observed differences in reactivity consistent with a substantial negative electrostatic potential that likely serves to attract and orient cationic ligands (*11*, *15*).

The Secondary Structure of TM4. The pattern of the residues that are accessible to the MTS reagents is consistent with TM4 being α-helical from W160$^{4.50}$ through C168$^{4.58}$ (Figure 5). The lack of an effect of the MTS reagents[2] on L174$^{4.64}$C suggests that the helix does not extend to this position, but a termination of the helix at 4.62 or 4.63 would be compatible with our results (see below). As discussed in

[2] If reaction of an MTS reagent with a substituted cysteine produces no change or only a small change in binding, then our use of binding as an indirect measure of reaction might produce a false negative conclusion about the accessibility of an engineered cysteine.

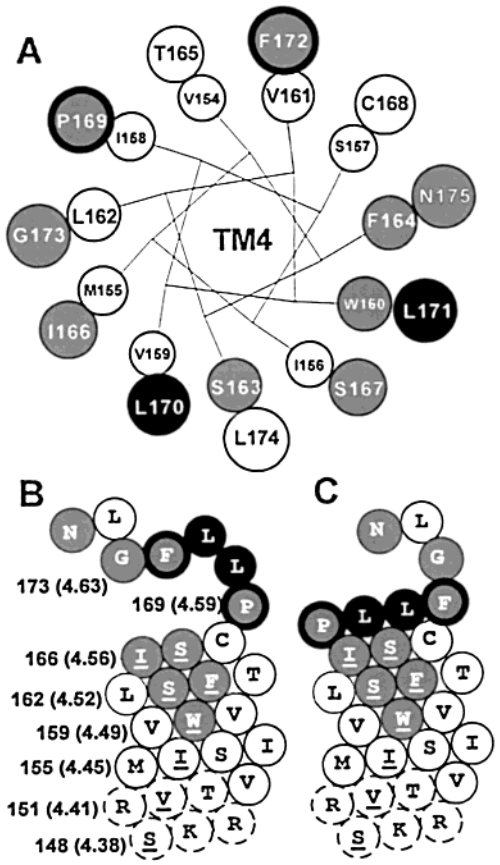

FIGURE 5: Helical wheel (A) and helical net (B, C) representations of the residues in and flanking the TM4 segment of the dopamine D2 receptor, summarizing conservation of residues and the effects of MTSEA on [^{3}H]*N*-methylspiperone binding. Reactive residues are solid filled for protected residues with rates of reaction ≥ 43 M^{-1} s^{-1}, gray for protected residues that were significantly inhibited by MTSEA but had a rate ≤ 9 M^{-1} s^{-1} and gray with bold borders for unprotected residues that were significantly inhibited by MTSEA. Open circles indicate that MTSEA had no significant effect on binding. In panel A, increasing sizes of the circles indicate increasingly extracellular localization of the indicated residues. In panels B and C, highly conserved residues in homologous GPCRs are underlined, and the dotted circles represent residues that were not mutated but are predicted to be in TM4.

Results, the presence of PP and PXP motifs suggests that the helix does not universally extend beyond 4.58. Thus, in the homologous GPCRs with PP or PXP motifs in this region, and perhaps in all homologous GPCRs, the secondary structure is most likely delimited as depicted in Figure 5B, in which P169$^{4.59}$ is shown as the first residue after the end of the α-helix. The secondary structure of the following loop is unknown, but a number of residues in this loop clearly are both accessible to reaction with the MTS reagents and protected by sulpiride, whereas the two residues P169$^{4.59}$ and F172$^{4.62}$ appear to be accessible but not protected.

An alternative representation of the secondary structure is shown in Figure 5, panel C, in which the helical structure continues through F172$^{4.62}$. This places L170$^{4.60}$ and L171$^{4.61}$, the sites that are highly reactive when substituted by Cys, in the most extracellular turn of the helix. Consequently, P169$^{4.59}$ and F172$^{4.62}$ are placed in the most extracellular turn where they could be accessible from the extracellular milieu but not directly face the binding-site crevice and thus not be protected by sulpiride.

It is not possible to assign the secondary structure of the D2 receptor definitively to either of these models, as they both are consistent with our data. The occurrence of PP and PXP motifs and the divergence of the residues at the extracellular end of TM4 in amine GPCRs make it very unlikely that the structure of all homologous receptors is a continuous α-helix through 4.62, even if all these receptors share a basic structural framework. It is not necessary, however, that the α-helical structure of a given TMS end at exactly the same position in every receptor. The D2 receptor may be representative of a subgroup of GPCRs that extends an additional helical turn at the extracellular end of TM4. Nonetheless, TM4 still appears to be the shortest helix in the bundle, consistent with the inferences from the rhodopsin low-resolution structure (*35*). Rhodopsin contains a PP motif at 4.59, consistent with a CT at 4.58, so that with an NT between 4.38 and 4.40, the length of TM4 is predicted to be only 18−20 residues.

The Extracellular Portion of TM4. As discussed in Results, the extracellular portion of TM4 from 4.59 to 4.68 is highly divergent among amine receptors (Table 4) yet subtype-selective in its conservation pattern. This finding is consistent with this region playing an important role in determining the functional similarity within subgroups of related proteins; it suggests that the structural divergence in the extracellular portion of TM4 is important for the discriminating functions of these receptors, such as ligand recognition and/or possibly ligand access to the binding-site crevice. The α-helical segment of TM4 is significantly shorter than other transmembrane helices, as shown in Figure 4, panel B, by comparison with the adjacent TM3 and TM5 helices. Because the high density of Arg/Lys residues within its cytoplasmic boundaries are expected to establish ionic interactions with negatively charged phospholipid headgroups, the cytoplasmic end of TM4 helix is likely to reside close to the cytoplasmic end of other TM helices with similar Arg/Lys residues. It follows that the portion of TM4 from 4.59 to 4.64, which is predicted to be nonhelical, resides in the transmembrane domain at the level of the ligand binding site, near D3.32 in TM3 and S5.43−S5.46 in TM5 (Figure 4B), which are identified contact sites for dopamine.

Indeed, mutagenesis studies in related GPCRs suggest that a number of residues in this region are critical determinants of agonist and/or antagonist binding, although it has been unclear which of these effects are direct and which are indirect. Consistent with the results presented here, mutation of P4.59 affected ligand affinity in a number of receptors including the muscarinic m3 receptors (*36*), C5a receptors (*37*), and thyrotropin receptors (*38*). Substitutions at 4.60 also decreased agonist and antagonist affinity in neurokinin type 1 and 2 receptors (*39*, *40*), angiotensin AT1 receptors (*41*), vasopressin V1a receptors (*42*), and dopamine D3 receptors (*43*). Several other residues between 4.61 and 4.72 have also been shown to contribute to agonist and/or antagonist binding either alone or in combination with other mutations (*39*, *40*, *43*−*47*) emphasizing the role of the extracellular portion of TM4 in ligand binding. In addition, Matsui et al. (*8*) showed with the use of photoaffinity labeling with agonists and antagonists that the peptide labeled by both ligands in the α2-adrenergic receptor corresponded to a fragment containing residues 4.44 through 4.66, suggesting that this portion of the receptor forms part of the ligand-binding domain.

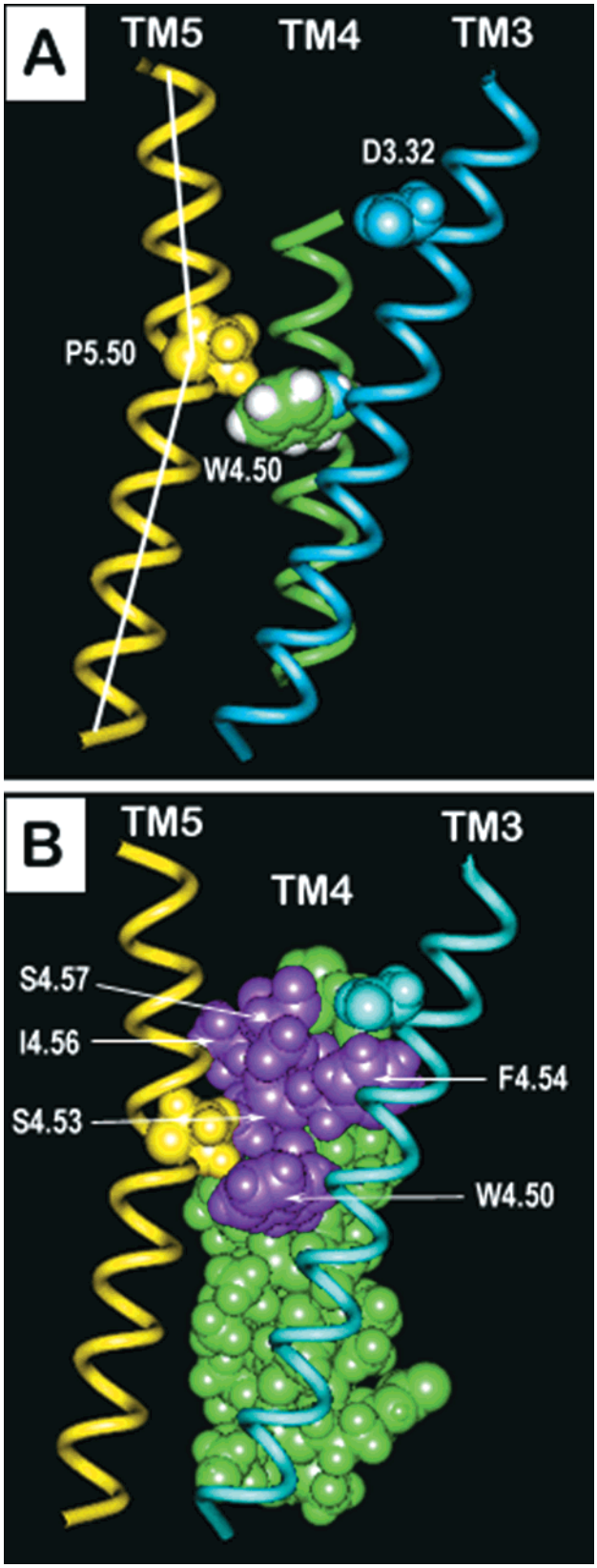

FIGURE 6: Molecular model of TM3, TM4, and TM5 in the dopamine D2 receptor. The top of the figure is the extracellular end and the bottom of the figure is the intracellular end of the TMSs. TM3 is shown in blue, TM4 is shown in green, and TM5 is shown in yellow. The accessible residues are shown in purple. (A) A ribbon representation shows the relative tilting and positions of these TMSs, including the Pro-kink of TM5, and the relative positions of W4.50, P5.50, and D3.32. (B) TM4 is shown in van der Waals representation, with the residues accessible in the binding-site crevice, between W4.50 and C4.58, shown in purple.

Curiously, none of the cysteine substitutions in TM4 of the dopamine D2 receptor significantly affected the binding affinity of the antagonist *N*-methylspiperone, whereas all the cysteine mutants between P169$^{4.59}$C and L174$^{4.64}$C had lower affinity for the antagonist sulpiride. This suggests that the two compounds bind, at least in part, to different regions of the receptor and/or that the compounds are differentially sensitive to mutations that alter the conformation of the extracellular region of TM4. We have observed a similar greater sensitivity of sulpiride to mutations at a number of positions in other TMSs (*12*−*15*, *48*).

The Tertiary Structure of TM4. The bulky Trp side chain at 4.50 represents a protuberance on the van der Waals surface of TM4 (Figure 6B) that faces the interior of the transmembrane bundle, which in turn requires a significant opening at the TM3−TM5 interface in the packing of the bundle. The opening is achieved by the kink induced in TM5 by P5.50 (*24*), which bends TM5 away from W4.50 (Figure 6). Thus, positioning of the bulky and highly conserved W4.50 in the model near the region where TM4 becomes accessible to the binding site crevice accommodates all the packing constraints and can satisfactorily explain the pattern of accessibility identified by SCAM, shown in purple in Figure 6B.

The secondary structure prediction described in Results is consonant with results from spin labeling studies in rhodopsin (*27*) and suggests that approximately 10−12 residues preceding W4.50 are in TM4 (Figure 5). The accessibility to MTS reagents, however, appears limited to the extracellular half of TM4, with the most cytoplasmic-accessible residue W160$^{4.50}$ predicted to lie roughly in the middle of the TMS (Figure 5). This is consistent with the low-resolution structure of rhodopsin viewed from the binding-site crevice looking outward, which, in the more intracellular portion of the receptor, places TM4 behind TM3 and TM5 where it becomes inaccessible from the binding-site crevice (*4*, *5*) (Figure 6). In contrast, the more extracellular portion of TM4 presents a face accessible to the binding-site crevice, and our results suggest that the bulky and highly conserved W160$^{4.50}$ is positioned near the point where TM4 becomes accessible.

The opening between TM3 and TM5 is most likely caused by the kink in TM5 induced by P5.50, resulting in the extracellular ends of TM3 and TM5 tilting significantly away from each other. The conserved disulfide bridge between C3.25 in TM3 and C5.31 in TM5 (Figure 4B) likely functions to stabilize the resulting structure by linking these two segments at an appropriate distance. This disulfide bridge has been shown to be essential for the stability of a number of GPCRs (*49*, *50*). Notably, the melanocortin and cannabinoid receptors, which lack the disulfide bridge, also lack the conserved P5.50, suggesting that the disulfide may not be essential for stability in the absence of the kink that bends TM5 away from TM3.

In summary, the present combination of results from SCAM experiments on TM4 with the analysis of sequence alignments and structure−function probing of cognate GPCR families place the current receptor model in the general context of the emerging structural understanding of GPCRs. Moreover, the putative nonhelical structure and the special sequence conservation properties of the extracellular portion of TM4 identified in this context suggest an important role for this region in ligand recognition.

NOTE ADDED IN PROOF

Our SCAM results and our proposal for the general orientation and location of TM4 with the TM bundle are consistent with the recent groundbreaking high resolution structure of the homologous GPCR rhodopsin [Palczewski et al. (2000) *Science 289*, 739]. Moreover, the extracellular end of TM4 in rhodopsin reveals a single turn of distorted kinked helix following Pro4.59 and Pro4.60, also consistent with our data and interpretation. The unique sequence properties we noted in GPCRs from 4.59 to 4.68 are noteworthy given that in rhodopsin these residues connect the top of TM4 with the beta strand structure that extends down into the binding-site crevice in rhodopsin where it interacts with retinal. Our results indicate that, like in rhodopsin, in the dopamine D2 receptor this loop may interact with the binding site and/or bound ligand. Curiously, the side chain of Trp4.50 in the rhodopsin structure does not protrude between TM3 and TM5 as we inferred but rather extends into the lipid behind TM3, but the accessibility data remain consistent with the rhodopsin structure, suggesting a structural congruity in this region.

ACKNOWLEDGMENT

We thank Drs. Olivier Civelli, Brian Kobilka, and Steve Rees for the human D2 receptor cDNA, epitope-tagged β2 adrenergic receptor cDNA, and the pcin4 vector, respectively. We thank Thomas Livelli for the HEK 293 and the 293-TSA cells and for valuable advice. We thank Myles Akabas and Arthur Karlin for valuable discussion and for comments on this manuscript.

REFERENCES

1. Civelli, O., Bunzow, J. R., Grandy, D. K., Zhou, Q. Y., and Van Tol, H. H. (1991) *Eur. J. Pharmacol. 207*, 277−86.
2. Strader, C. D., Fong, T. M., Tota, M. R., Underwood, D., and Dixon, R. A. (1994) *Annu. Rev. Biochem. 63*, 101−32.
3. Schertler, G. F., Villa, C., and Henderson, R. (1993) *Nature 362*, 770−2.
4. Unger, V. M., Hargrave, P. A., Baldwin, J. M., and Schertler, G. F. (1997) *Nature 389*, 203−6.
5. Baldwin, J. M., Schertler, G. F., and Unger, V. M. (1997) *J. Mol. Biol. 272*, 144−64.
6. Ballesteros, J. A., and Weinstein, H. (1995) *Methods Neurosci. 25*, 366−428.
7. van Rhee, A. M., and Jacobson, K. A. (1996) *Drug Dev. Res. 37*, 1−38.
8. Matsui, H., Lefkowitz, R. J., Caron, M. G., and Regan, J. W. (1989) *Biochemistry 28*, 4125−30.
9. Akabas, M. H., Stauffer, D. A., Xu, M., and Karlin, A. (1992) *Science 258*, 307−10.
10. Javitch, J. A., Li, X., Kaback, J., and Karlin, A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 10355−9.
11. Javitch, J. A., Fu, D., Chen, J., and Karlin, A. (1995) *Neuron 14*, 825−31.
12. Javitch, J. A., Fu, D., and Chen, J. (1995) *Biochemistry 34*, 16433−9.
13. Fu, D., Ballesteros, J. A., Weinstein, H., Chen, J., and Javitch, J. A. (1996) *Biochemistry 35*, 11278−85.
14. Javitch, J. A., Ballesteros, J. A., Weinstein, H., and Chen, J. (1998) *Biochemistry 37*, 998−1006.
15. Javitch, J. A., Ballesteros, J. A., Chen, J., Chiappa, V., and Simpson, M. M. (1999) *Biochemistry 38*, 7961−7968.
16. Stauffer, D. A., and Karlin, A. (1994) *Biochemistry 33*, 6840−9.
17. Rees, S., Coote, J., Stables, J., Goodson, S., Harris, S., and Lee, M. G. (1996) *BioTechniques 20*, 102−110.
18. Sander, C., and Schneider, R. (1991) *Proteins 9*, 56−68.
19. Horn, F., Weare, J., Beukers, M. W., Horsch, S., Bairoch, A., Chen, W., Edvardsen, O., Campagne, F., and Vriend, G. (1998) *Nucleic Acids Res. 26*, 275−9.
20. Vriend, G. (1990) *J. Mol. Graph. 8*, 52−56.
21. Ballesteros, J., Kitanovic, S., Guarnieri, F., Davies, P., Fromme, B. J., Konvicka, K., Chi, L., Millar, R. P., Davidson, J. S., Weinstein, H., and Sealfon, S. C. (1998) *J. Biol. Chem. 273*, 10445−53.
22. Javitch, J. A., Fu, D., Liapakis, G., and Chen, J. (1997) *J. Biol. Chem. 272*, 18546−9.
23. Ballesteros, J. A., and Weinstein, H. (1992) *Biophys. J. 62*, 107−9.
24. Ballesteros, J. A., and Weinstein, H. (1995) *Biophys. J. 68*, A446.
25. Elling, C. E., Thirstrup, K., Nielsen, S. M., Hjorth, S. A., and Schwartz, T. W. (1997) *Ann. N. Y. Acad. Sci. 814*, 142−51.
26. Sealfon, S. C., Chi, L., Ebersole, B. J., Rodic, V., Zhang, D., Ballesteros, J. A., and Weinstein, H. (1995) *J. Biol. Chem. 270*, 16683−8.
27. Farahbakhsh, Z. T., Ridge, K. D., Khorana, H. G., and Hubbell, W. L. (1995) *Biochemistry 34*, 8812−9.
28. Komiya, H., Yeates, T. O., Rees, D. C., Allen, J. P., and Feher, G. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 9012−6.
29. Donnelly, D., Overington, J. P., Ruffle, S. V., Nugent, J. H., and Blundell, T. L. (1993) *Protein Sci. 2*, 55−70.
30. Donnelly, D., and Cogdell, R. J. (1993) *Protein Eng. 6*, 629−35.
31. Davison, M. D., and Findlay, J. B. (1986) *Biochem. J. 236*, 389−95.
32. MacArthur, M. W., and Thornton, J. M. (1991) *J. Mol. Biol. 218*, 397−412.
33. Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. E., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−42.
34. Roberts, D. D., Lewis, S. D., Ballou, D. P., Olson, S. T., and Shafer, J. A. (1986) *Biochemistry 25*, 5595−601.
35. Schertler, G. F. (1998) *Eye 12*, 504−10.
36. Wess, J., Nanavati, S., Vogel, Z., and Maggio, R. (1993) *EMBO J. 12*, 331−8.
37. Kolakowski, L. F., Jr., Lu, B., Gerard, C., and Gerard, N. P. (1995) *J. Biol. Chem. 270*, 18077−82.
38. Stein, S. A., Oates, E. L., Hall, C. R., Grumbles, R. M., Fernandez, L. M., Taylor, N. A., Puett, D., and Jin, S. (1994) *Mol. Endocrinol. 8*, 129−38.
39. Huang, R. R., Vicario, P. P., Strader, C. D., and Fong, T. M. (1995) *Biochemistry 34*, 10048−55.
40. Fong, T. M., Yu, H., Cascieri, M. A., Underwood, D., Swain, C. J., and Strader, C. D. (1994) *J. Biol. Chem. 269*, 14957−61.
41. Perlman, S., Costa-Neto, C. M., Miyakawa, A. A., Schambye, H. T., Hjorth, S. A., Paiva, A. C., Rivero, R. A., Greenlee, W. J., and Schwartz, T. W. (1997) *Mol. Pharmacol. 51*, 301−11.
42. Mouillac, B., Chini, B., Balestre, M. N., Elands, J., Trumpp-Kallmeyer, S., Hoflack, J., Hibert, M., Jard, S., and Barberis, C. (1995) *J. Biol. Chem. 270*, 25771−7.
43. Alberts, G. L., Pregenzer, J. F., and Im, W. B. (1998) *Mol. Pharmacol. 54*, 379−88.
44. Pan, Y., Wilson, P., and Gitschier, J. (1994) *J. Biol. Chem. 269*, 31933−7.
45. Hjorth, S. A., Schambye, H. T., Greenlee, W. J., and Schwartz, T. W. (1994) *J. Biol. Chem. 269*, 30953−9.
46. Werge, T. M. (1994) *J. Biol. Chem. 269*, 22054−8.
47. Green, S. A., Cole, G., Jacinto, M., Innis, M., and Liggett, S. B. (1993) *J. Biol. Chem. 268*, 23116−21.
48. Javitch, J. A., Fu, D., and Chen, J. (1996) *Mol. Pharmacol. 49*, 692−8.
49. Noda, K., Saad, Y., Graham, R. M., and Karnik, S. S. (1994) *J. Biol. Chem. 269*, 6743−52.
50. Karnik, S. S., Sakmar, T. P., Chen, H. B., and Khorana, H. G. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8459−63.
51. Cheng, Y., and Prusoff, W. H. (1973) *Biochem. Pharmacol. 22*, 3099−108.

BI001069M